

Edited by Martin Engebretsen and Helen Kennedy

# Data Visualization in Society

Amsterdam University Press

Data Visualization in Society

## Data Visualization in Society

Edited by Martin Engebretsen and Helen Kennedy

Amsterdam University Press

The publication of this book is made possible by a grant from the Research Council of Norway (grant number 25936).

Cover illustration: Prisca Schmarsow of Eyedea Studio

Cover design: Coördesign, Leiden Lay-out: Crius Group, Hulshout

 ISBN
 978 94 6372 290 2

 e-ISBN
 978 90 4854 313 7

 DOI
 10.5117/9789463722902

 NUR
 811



Creative Commons License CC BY NC ND (http://creativecommons.org/licenses/by-nc-nd/3.o)

© All authors / Amsterdam University Press B.V., Amsterdam 2020

Some rights reserved. Without limiting the rights under copyright reserved above, any part of this book may be reproduced, stored in or introduced into a retrieval system, or transmitted, in any form or by any means (electronic, mechanical, photocopying, recording or otherwise).

Every effort has been made to obtain permission to use all copyrighted illustrations reproduced in this book. Nonetheless, whosoever believes to have rights to this material is advised to contact the publisher.

## **Table of Contents**

List of tables	
List of figures	
Acknowledgements	
Foreword: The dawn of a philosophy of visualization	17
Alberto Cairo, Knight Chair at the University of Miami and author of How Charts Lie	
1. Introduction: The relationships between graphs, charts, maps	
and meanings, feelings, engagements	19
Helen Kennedy and Martin Engebretsen	
Section I Framing data visualization	
2. Ways of knowing with data visualizations <i>Jill Walker Rettberg</i>	35
3. Inventorizing, situating, transforming: Social semiotics and data visualization <i>Giorgia Aiello</i>	49
4. The political significance of data visualization: Four key perspectives	63
Torgeir Uberg Nærland Section II Living and working with data visualization	

5.	5. Rain on your radar: Engaging with weather data	
	visualizations as part of everyday routines	77
	Eef Masson and Karin van Es	
6.	Between automation and interpretation: Using data	
	visualization in social media analytics companies	95
	Salla-Maaria Laaksonen and Juho Pääkkönen	

7.	Accessibility of data visualizations: An overview of European statistics institutes <i>Mikael Snaprud and Andrea Velazquez</i>	111
8.	Evaluating data visualization: Broadening the measurements of success <i>Arran L. Ridley and Christopher Birchall</i>	127
9.	Approaching data visualizations as interfaces: An empirical demonstration of how data are imag(in)ed <i>Daniela van Geenen and Maranke Wieringa</i>	141
10.	Visualizing data: A lived experience Jill Simpson	157
11.	Data visualization and transparency in the news Helen Kennedy, Wibke Weber, and Martin Engebretsen	169

## Section III Data visualization, learning, and literacy

12.	What is visual-numeric literacy, and how does it work? <i>Elise Seip Tønnessen</i>	189
13.	Data visualization literacy: A feminist starting point Catherine D'Ignazio and Rahul Bhargava	207
14.	Is literacy what we need in an unequal data society? <i>Lulu Pinney</i>	223
15.	Multimodal academic argument in data visualization Arlene Archer and Travis Noakes	239
Sec	tion IV Data visualization semiotics and aesthetics	
16.	What we talk about when we talk about beautiful data visualizations <i>Sara Brinch</i>	259

17.	A multimodal perspective on data visualization <i>Tuomo Hiippala</i>	277
18.	Exploring narrativity in data visualization in journalism <i>Wibke Weber</i>	295
19.	The data epic: Visualization practices for narrating life and death at a distance <i>Jonathan Gray</i>	313
20.	What a line can say: Investigating the semiotic potential of the connecting line in data visualizations <i>Verena Elisabeth Lechner</i>	329
21.	Humanizing data through 'data comics': An introduction to graphic medicine and graphic social science Aria Alamalhodaei, Alexandra Alberda, and Anna Feigenbaum	347
See	ction V Data visualization and inequalities	
22.	Visualizing diversity: Data deficiencies and semiotic strategies John P. Wihbey, Sarah J. Jackson, Pedro M. Cruz, and Brooke Fou- cault Welles	369
23.	What is at stake in data visualization? A feminist critique of the rhetorical power of data visualizations in the media <i>Rosemary Lucy Hill</i>	391
24.	The power of visualization choices: Different images of patterns in space <i>Britta Ricker, Menno-Jan Kraak, and Yuri Engelhardt</i>	407
25.	Making visible politically masked risks: Inspecting unconventional data visualization of the Southeast Asian haze <i>Anna Berti Suman</i>	425
26.	How interactive maps mobilize people in geoactivism <i>Miren Gutiérrez</i>	441

Index

457

## List of tables

Table 7.1	Overview of NSI websites and accessibility score from	
	the WTKollen checker tool	117
Table 18.1	Telling vs. showing	298
Table 18.2	Narrative constituents	299
Table 19.1	Comparison of features contributing to aesthetics of	
	data sublime in Halloran 2015 and 2018	322
Table 19.2	Comparison of features contributing to connection	
	between scale and individual in Halloran 2015 and 2018	323
Table 26.1	Three maps seen from DeSoto's and Muehlenhaus's	
	categorizations	451

## List of figures

Figure 2.1	The height of Belgians from 18 to 20 years	41
Figures 5.1 and 5.2	Default views for the Buienradar website and	
	(Android) app, for Monday April 30, 2018 at	
	around 11:35 a.m. CET	79
Figure 5.3	Weather report with textual and graphic ele-	
	ments in NRC Handelsblad (a Dutch national	
	'quality' newspaper) for the weekend of 21 and	
	April 22, 2018	83
Figure 5.4	Shower radar and rain graph visualizations on	
	the Buienradar website, set to Amsterdam, for	
	Monday April 30, 2018, 11:35 a.m. CET	85
Figure 9.1	'Raw' version of the network graph in the	
	'Overview' after the data import into Gephi	147
Figure 9.2	Spatialized graph after the application of	
	ForceAtlas 2 (Scaling: 2.0, Gravity: 20.0; node	
	size based on degree)	147
Figure 9.3	Exported, spatialized, filtered, and annotated	
	graph	149
Figure 10.1	Visualizing mental illness: A day of OCD	161
Figure 12.1	Screenshot of Google Public Data	193
Figure 12.2	Screenshot of the starting image on Gap-	
	minder tools	193
Figure 12.3	Screenshot of group 7's screen after placing	
	the word document side by side with the same	
	graph as the one displayed in Figure 12.1	197
Figure 12.4	Versions of graph to answer task 2 on child	
	mortality in three countries. a) Group 2 and 3	
	with intended axis variables b) Group 1, c)	
	Group 4, d) Group 5	201
Figure 13.1	Some of the sketches students created	211
Figure 13.2	The data mural	212
Figure 13.3	The WTFcsv results screen	213
Figure 15.1a	Visual data hedging through the use of a	
	confidence interval	244
Figure 15.1b	An alternate visual form of hedging with	
	maximum y-axis	244

Figure 15.1c	Another visual form of hedging with the	
	maximum and minimum values labelled	244
Figure 15.2	Maximum and minimum values indicated	
	using separate line graphs	245
Figure 15.3	Nyanga versus Newlands	247
Figure 15.4	Language, Education and Internet Access in	
-	neighbouring wards of Cape Town	250
Figure 16.1	Charles Joseph Minard's map of Napoleon's	
C	Russian Campaign 1812-1813	265
Figure 16.2	Fernanda Viégas and Martin Wattenberg's	
	Wind Map	267
Figure 16.3	Charlie Clark's Blade Runner from the project	
-	'The Color of Motion'	268
Figure 16.4	Valentina D'Efilippo's Poppy Field	271
Figure 16.5a and b	Front and backside of Week 8 (Phone Addic-	
C -	tion / A Week of Phone Addiction), Dear Data	272
Figure 17.1	A static information graphic reporting on the	
-	death of the last male northern white rhino	283
Figure 17.2	Four screenshots from a non-interactive	
-	dynamic data visualization 'Temperature	
	anomalies arranged by country 1900–2016'	
	showing temperature anomalies arranged by	
	country 1900–2016	285
Figure 17.3	The Seas of Plastic, an interactive dynamic	
C	data visualization	288
Figure 17.4	The decomposition of (1) static information	
-	graphics, (2) non-interactive data visualiza-	
	tions, and (3) interactive data visualizations	
	into canvases	289
Figure 18.1	Telling, showing, telling. A modified version of	
-	the Martini-Glass structure	300
Figure 18.2	Screenshot of the intro of the data visualiza-	
	tion '20 years, 20 titles'	302
Figure 18.3	Sequential pattern with scrolling and zooming	
-	out. Drawn after the graphic 'Mass exodus: the	
	scale of the Rohingya crisis'	304
Figure 19.1	The white timelines of individual lives ending	
	in the red block of WWII	318
Figure 19.2	Group of silhouettes rendered equivalent to an	
	isotype figure	318

Figure 19.3	Panning up a long column of Soviet deaths	319
Figure 19.4	Comparing population of total living with	
	total dead	320
Figure 19.5	Visualizing nuclear disarmament alongside	
	proliferation	321
Figure 20.1	Example of a data visualization using lines to	
	represent the connections between sanitary	
	problems (central group of purple letter and	
	number codes) and the restaurants in Man-	
	hattan, NYC they occurred in, represented as	
	dots in the outer circle	331
Figure 20.2	Three exemplary compositions of connecting	
	lines and their connected elements	339
Figure 20.3	Visualization of the connections related to	
	Sigmundur Davíð Gunnlaugsson	340
Figure 20.4	Two visualizations of spatial movement,	
	using a top-down angle in a route map (upper	
	picture) and a frontal perspective in an arc	
	diagram (lower picture)	342
Figure 21.1	Hand-drawn amfAR line graph	354
Figure 21.2	HIV virus cell	354
Figure 21.3	ATZ pills prescription	355
Figure 21.4	HAART medication introduction	356
Figure 21.5	Explanation of methane gas	358
Figure 21.6	Explanation of the Gulf Stream	359
Figure 21.7	Graph of the Earth's surface temperature from	
	year 1000-2100	359
Figure 22.1	Map/chart included in 1896 US Census	
	documents, showing growth of racial and	
	demographic groups and territorial expan-	
	sion	372
Figure 22.2	Chart included in 1983 US Census Bureau	
	documents, showing the relative contribution	
	of various continents to immigration totals in	
	the United States	373
Figure 22.3	Historical rendering by sociologist W. E. B. Du	
	Bois of trajectory of African slave trade to the	
	Americas	374
Figure 22.4	Native American map rendering on deerskin	
	of tribal information and location	375

Figure 22.5	Visualization of US immigration as metaphori-	
-	cal rings in a growing tree trunk, with each	
	dash representing 100 immigrants and each	
	ring representing one decade. The image is	
	based on Census data relating to persons'	
	origin at birth, 1830-2015	381
Figure 22.6	Evolution over time of visual simulation of	
0	US immigration as metaphorical rings in a	
	growing tree trunk	382
Figure 22.7	Examples of visual patterns in specific US	-
0 .	states relating to immigration as metaphorical	
	rings in a growing tree trunk. White cells	
	represent native-born persons, while coloured	
	cells represent immigrants	384
Figure 23.1	Most Americans say they don't know enough	
0	about the abortion pill to say if it is safe and	
	effective	395
Figure 23.2	Abortion Rate & Ratio vs. Poverty Rate	397
Figure 23.3	Abortion in the United States	401
Figure 24.1	The traditional cartographic workflow with	
	special attention to feminist cartography. The	
	diagram shows the steps needed to go from	
	data to map or diagram, split between the	
	cartographic data analysis and the carto-	
	graphic design. Parallel to this workflow is the	
	feminist data visualization workflow	413
Figure 24.2	The Gender Inequality Index: a) a table with	
	the index value for each country; b) a bar	
	graph with the index order from high to low	
	inequality; c) a map with the geographic	
	distribution of inequality emphasizing high	
	inequality; d) the variables included in the	
	Gender Inequality Index	415
Figure 24.3	Choosing colours and colour ramps for a	
	choropleth map: a) red, emphasizing high	
	inequality; b) red, emphasizing low inequality;	
	c) green, emphasizing high inequality; d)	
	green, emphasizing low inequality.	417

Figure 24.4	The influence of the map projection: a) the	
	equal-area Mollweide projection; b) the	
	conformal Mercator projection; c) a cartogram	
	based on the country population	419
Figure 25.1	The Global Forest Watch Fires Map by the	
	World Resource Institute	431
Figure 25.2	Kepo Hutan Map by Greenpeace in collabora-	
	tion with the World Resource Institute	432
Figure 26.1	A moment in the 15M map	442
Figure 26.2	A moment of 'Left-to-die boat' map	445
Figure 26.3	Sierra Loba, as it engages in irregular opera-	
	tions in Senegalese coastal waters	449

## Acknowledgements

We would like to thank the Research Council of Norway for funding INDVIL, a research project from which this edited collection emerged (Innovative Data Visualization and Visual-Numeric Literacy (https://indvil.org, grant number 259536). We thank our colleagues on INDVIL—Sara Brinch, Verena Lechner, Torgeir Uberg Nærland, Jill Walker Rettberg, Mikael Snaprud, Elise Seip Tønnessen, and Wibke Weber—for a stimulating and productive three-year collaboration. They are all authors of chapters in this book. We are grateful to all of the authors who have contributed to this collection, for their excellent contributions, their receptiveness to feedback and their patience during the long process of bringing the book to fruition. We are extremely grateful to the University of Agder for hosting the INDVIL project and for the administrative support provided to us. In June 2018 we held a seminar with many of the book's authors at the Metochi Study Centre on the island of Lesbos. We thank the staff at Metochi for the wonderful service they provided and the fantastic atmosphere which enabled us to work hard together on the book. We thank reviewers for taking the time to read our many chapters with care and for providing helpful feedback. We are grateful also to Alberto Cairo for his extremely generous foreword, and to esteemed data visualization scholars for endorsing this volume on its cover. We want to express our enormous gratitude to all of the brilliant data visualizers whose work inspires our own, especially Andy Kirk, for ongoing support and openness to dialogue. We could not have made it through the final stages of producing this book without the invaluable help of Robin Steedman, whose work as an editorial assistant has been faultless. Thank you Robin. We also thank Prisca Schmarsow of Eyedea Studio (https://eyedeastudio.london/) for lovely cover design and the staff at the Amsterdam University Press for their assistance along the way. Finally, Helen would like to thank Martin, and Martin would like to thank Helen, for the enjoyable and rewarding experience of editing this book together.

## Foreword: The dawn of a philosophy of visualization

Alberto Cairo, Knight Chair at the University of Miami and author of How Charts Lie

Geographer John Pickles once wrote that 'GIS is a set of tools, technologies, approaches and ideas that are vitally embedded in broader transformations of science, society, and culture'. That's true of data visualization too, therefore the relevance of the book that you have in your hands, *Data Visualization in Society*.

I often joke—although I'm inclined to believe—that a field X reaches maturity when a parallel field of 'philosophy of X' springs into existence. That hasn't happened yet with data visualization, at least formally. Might we be on the path to it, though? I hope so. Some books have paved the way. Think of David J. Staley's *Computers, Visualization, and History*, Charles Kostelnick and Michael Hassett's *Shaping Information*, and Wolff-Michael Roth's *Toward an Anthropology of Graphing*, all from the early 2000s. Or, more recently, Orit Halpern's *Beautiful Data* (2014), Johanna Drucker's *Graphesis* (2014), R. J. Andrews's *Info We Trust* (2019), Sandra Rendgen and Julius Wiedemann's *History of Information Graphics* (2019), or the upcoming *Data Feminism* (2020), by Catherine D'Ignazio and Lauren Klein, who have also contributed to this volume.

Books like these prove that writing about visualization doesn't mean just thinking about how to *design* visualizations, but also about what visualization *is*, why it is the way it is—and what it *could* be. Data visualization is a technology—or set of technologies—and, like artefacts such as the clock, the compass, the abacus, or the map, it transforms the way we see and relate to reality. As Langdon Winner suggested in *The Whale and the Reactor* (1986), a foundational book in the phenomenological philosophy of technology, to create technologies doesn't consist just of crafting stuff; rather, when technologies come about 'new worlds are being made'. What 'new worlds' does visualization generate? That's a question for a potential philosophy of visualization.

A philosophy of visualization may derive themes, methodologies, and language from a wide range of disciplines: epistemology, sociology, semiotics, history, ethics, critical theory fields such as critical cartography, or from the philosophies of science, statistics, art, and—perhaps more strongly than any

other—the philosophy of technology. Philosophers of visualization should reason about visualization's history, assumptions, conventions, practices, and impacts on individuals, cultures, and societies. They will combine the observational, descriptive, and hermeneutical—dealing with what currently exists and why—the normative—, asking what should or shouldn't exist or happen—and the critical—, challenging visualization's core tenets.

Data Visualization in Society is a collection of chapters by scholars and professionals who don't call themselves philosophers of visualization but who, in practice, operate as such. I see this book as a relevant step toward the possible inception of the philosophy of data visualization as a discipline. I hope it will serve as a starting point for many inquiries by other thinkers. This includes myself: I read all chapters with pleasure and took copious notes on the margins. I know these scribbles will later echo in my own work.

That's the virtue of the best philosophical writing: it doesn't aspire to settle matters outright, but to inspire further reflection. *Data Visualization in Society* may spur questions such as: Does visualization pretend to be 'objective', or is it just wrongly perceived as such? What does 'objective' mean in the first place? What is the influence of visualization on politics? Is numeracy—numerical literacy—enough to design or read visualizations? Doesn't the fact that a substantial portion of the public isn't numerate—or 'graphicate'—deepen existing inequalities and even create new ones? What do we mean when we say that a visualization is 'beautiful'? Is the goal of visualization to convey facts and data, or can it also spark profound emotional experiences? If so, how? And many more.

The variety of topics and approaches of the chapters in this book is astounding, but what most have in common is an open ending: they are links in a chain of reasoning—a dialogue—that extends from the distant past and that, conceivably, and with the contribution of a large critical mass of academics and practitioners of the craft, will continue beyond the foreseeable future. That's where you come in: does any of these chapters inspire you? Do you agree or disagree with it? Reason why. Argue. Establish a conversation with it. Write and publish, and be open to further responses and critiques. That's how philosophy begins.

## 1. Introduction: The relationships between graphs, charts, maps and meanings, feelings, engagements

Helen Kennedy and Martin Engebretsen

Today we are witnessing an increased use of data visualization in a range of domains and genres. In journalism, education, and public information as well as in workplaces, diverse forms of graphs, charts, and maps are used to explain, persuade, and tell stories. At best, visual representations of statistics and other, often quantitative data can convey complex facts and patterns quickly and effectively. At worst, they can appear confusing or manipulative. In an era in which more and more data are produced and circulated through online networks, and digital tools make visualization production increasingly accessible, it is important to study the conditions under which such visual texts are generated, disseminated and thought to benefit processes of sense-making, learning, and engaging.

Data visualization is not new. The graphical representation of numeric information has roots in early map-making, and grew in importance with the widespread use of data and statistics for planning and commerce in the nineteenth century (Friendly, 2008). Still, in our contemporary society, several factors contribute to give data visualization a social relevance on a scale we have not seen before. One of these factors, as Kennedy, Hill, Aiello, and Allen (2016b, p. 715) put it, is that '[...] data are becoming increasingly valued and relied upon, as they come to play an ever more important role in decision-making and knowledge about the world'.

In other words, more data are generated, gathered, stored, and made accessible than ever before. Data gathering takes place in many domains, often by law, including commerce, education, health, transport, and cultural and social life. These data offer insights into societal patterns otherwise invisible and unnoticed. Such documentation has been conducted for decades, but technological and other developments have led to its sharp

Engebretsen, M. and H. Kennedy (eds.), *Data Visualization in Society*. Amsterdam: Amsterdam University Press, 2020 DOI 10.5117/9789463722902\_CH01

increase, and data are now being gathered in huge volumes as a result of new techniques of measurement. These combined phenomena, sometimes called 'datafication' (Mayer-Schönberger & Cukier, 2013, p. 78) are understood as a transformation disrupting the social world in all its forms (Couldry, 2016).

Furthermore, to make data accessible to publics, rather than remaining a useful source only for experts and decision-makers, a range of actors have campaigned to open up public data, to make them reusable for a variety of activities and democratic purposes. Open data initiatives and related campaigning activities contribute to accelerate the spread of data visualization, which often serve as a main entry point to data for non-experts.

Another important driver in the spread of data visualizations is the development of related technology. New tools and techniques for harvesting, filtering, analysing, and visualizing data make these processes easier and cheaper. We are also witnessing new arenas for dissemination of and engagement with data, as data-based techniques are increasingly used for informative, persuasive, and rhetorical purposes in political campaigns, health communication, education, and in newsrooms, where new data visualization teams are being constructed, combining visual creativity with data science skills and other domain expertise (Engebretsen, Kennedy, & Weber, 2018).

As a result of these varied processes, data visualizations have innovative semiotic forms and result in novel types of communication and interactivity. This implies that their potential for meaning-making, for evoking emotions, democratic participation, and other forms of engagement is also in a state of transformation. So, while data visualizations have a growing importance in society, their novel forms and uses mean that our understanding of how they work as semiotic and aesthetic phenomena and how they support or hinder personal and social agency is also in flux. These transformations do 'good' or 'bad': do they promote understanding and engagement, as some commentators argue (e.g. Few, 2008 and Cairo, 2013), or do they do ideological work, privileging certain views of the world, as others claim (e.g. Barnhurst, 1994 and Latour, 1986; sources taken from Kennedy et al., 2016b)?

The phenomena described here, both the new uses of data visualization and the debates about them, form the focus of this book, which draws on a range of research and development projects to reflect on data visualization in society. The book addresses these questions:

- Where and how do citizens and publics engage with data visualizations, and for what reasons?
- In what new social and cultural contexts are data visualizations emerging, and to what ends?

- How do data visualizations create meanings in the various social and cultural arenas in which they appear, and what are their discursive roles and functions?
- How do data visualizations arouse feelings in their audiences, what kinds of emotional responses are activated, and to what ends?
- What does literacy mean when it comes to data visualization, and how can data visualization literacy be enhanced?
- What kinds of aesthetic characteristics do data visualizations have?
- What is the political significance of data visualization, and in what ways do data visualizations play a role in citizens' participation in democratic systems?

#### What do we mean by data and data visualization?

In a scientific context, data are generally understood to result from the generation, collection, observation, or registration of objects, events, or processes suitable to serve some analytical purpose. Similarly, in the context of data visualization, data can be anything that can be subjected to categorization, abstraction, and translation into graphical representation: persons, places, documents, relations, sentences, salaries, to mention some examples. A main distinction is between qualitative data and quantitative data. While qualitative data are valued for the uniqueness of each individual unit, be it a poem, a sentence, or an interview, quantitative data are valued for characteristics shared by all or many units in a dataset. It is their shared characteristics that make them objects for counting or measuring, and thus for numeric representation and statistical processing.

Both qualitative and quantitative data can be visualized. It is possible to visualize semantic structures in a novel, or networks of relationships between the works in an art collection, as seen, for example, in the work of Stefanie Posavec (http://stefanieposavec.com/). Most, but not all, of the contributions in this book focus on the visualization of quantitative data, for the reasons given above—that is, because their proliferation and increasing openness, and the enhanced availability of related tools, make them a socially and culturally significant phenomenon.

Numeric data can be structured or unstructured. Structured data have been subjected to statistical treatment and are typically represented as numbers in a table, with columns and rows presenting units and variables and numeric values positioned in cells. A common example is the datasets accessible from national statistics institutes (NSIs) which are often presented

to the public in tabular form. Unstructured data have not been subjected to any statistical or structuring processing, and appear as 'raw' data in an analogue or digital register, until the data are structured by someone with an intention to use them for some specific purpose. An automatic registration of cars driving through a tollbooth is one example.

'Big data' is a fashionable concept, although its use is rarely accompanied by a shared understanding of what it means or how it differs from 'small' data. Big data have been said to be characterized by three Vs: volume, variety, and velocity. More recently, additional Vs have been proposed, such as variability and value (http://whatis.techtarget.com/definition/3Vs; see also Kitchin, 2014 for additions which don't begin with V). When we talk about datasets consisting of thousands of rows of data, or new streams of data created every second, we are talking about big data. Data harvested from a social media platform, or from the activities on the finance market, are some examples. But exactly when data become big is hard to define.

In the same way that it is hard to distinguish between big and other data, the differences between data visualization, information visualization, information graphics, and scientific visualization are also blurred. As Kennedy and Allen write, data visualization 'has data at its heart' (2016, p. 309), and it often uses abstract, geometrical forms to represent numeric values and relations. In contrast, an information graphic explains phenomena graphically but may contain no numeric data, or it presents data in charts alongside other illustrations, like photographs or drawings. Scientific visualization is a concept mostly used in highly specialized, expert-to-expert contexts, for example within medicine and biology. Here, visualizations are used to illuminate specific aspects of certain physical objects or processes, and may include simulations, drawings, or processing of magnetic resonance imaging (Ambrosio, 2015).

Data visualizations are a discursive resource used in the dissemination of statistical information and often numeric data. In this book, data visualizations are understood as graphical representations of data which are primarily, but not solely, numeric. What's more, they are abstractions and reductions of the world, the result of human choices, social conventions, and technological processes and affordances, relating to generating, filtering, analysing, selecting, visualizing, and presenting data. Data visualizations (also called *dataviz* or DV) are created to 'facilitate understanding', to use Kirk's term (2016, p. 19; see also Borgo et al., 2013; Cairo, 2013), but they can also facilitate other things, such as persuasion. Consequently, we understand data visualizations as cultural artefacts with distinct semiotic, aesthetic, and social affordances. There is, however, much more to data visualization than what can be captured in any simple definition, as will become evident throughout the book.

#### How can dataviz produce meanings, feelings, and engagements?

In this book, we relate the social power of data visualizations to their abilities to produce meanings, feelings, and engagements in their users and audiences. Processes of socially situated meaning-making are best described in the field of social semiotics, first developed by the Australian linguist Michael Halliday (1978), later adapted to visual and multimodal artefacts by Gunther Kress and Theo van Leeuwen (1996) and others. In social semiotic theory, the meaning of semiotic material (which can include words, images, colours, and more) can be traced in three different dimensions, each relating to an aspect of the situation of communication. These are:

- 1. The field (or topic) of discourse. How does the semiotic material represent the world or ideas about the world? This is known as *ideational meaning*.
- 2. The participants involved in the process of communication. How does the semiotic material reflect, establish, or change the social relations between the participants? This is known as *interpersonal meaning*.
- 3. The semiotic resources activated in the process. How do all the elements of the semiotic material unite in a textual whole? This is known as *compositional meaning*.

In many situations, the semiotic material in question will be identified as 'a text', such as a multimodal webpage with words, images, and colours organized in a specific user interface. In other contexts, meaning is made through semiotic resources not conventionally identified as texts, such as buildings, clothes, and sculptures. Such artefacts nonetheless carry meaning based on certain culturally and historically formed conventions. The artefacts that this book is concerned with, data visualizations, will normally be produced, distributed, and used in ways comparable to other multimodal and mediated text types.

Semiotic interpretation and aesthetic experience (that is, our sensory impressions, as well as judgements based on taste) go hand in hand in our encounters with texts and other cultural artefacts, and where one stops and the other begins is hard to identify. Our encounters with form, colour, and composition are informed by bodily experience as well as aesthetic judgement, and so the aesthetic (as well as the semiotic) aspects of data visualization need to be taken into account. Also relevant to a discussion of meaning in data visualization is the issue of 'knowledge regimes', or epistemology. What aspects of reality are privileged in a semiotic text based on visualized, numeric data? What kinds of truth are foregrounded, and what knowledge, values, and attitudes result? Data visualizations may

seem to reflect reality in a more direct way than words because they are based on numbers, which seem trustworthy (Porter, 1995). But this does not mean that they are more true, in the sense that they offer a more objective representation of the world. This issue informs several contributions to this book.

Data visualizations thus create meanings through visual and other codes. But they also generate feelings, by which we mean the emotional responses that are connected to human encounters with data visualizations. Meanings and feelings are inseparable in our situated interactions with texts. They influence each other, and together they form our responses to the texts and artefacts with which we interact (Lemke, 2015). A recent study by Kennedy and Hill (2017) revealed that data visualizations awaken a wide range of feelings in people who engage with them, activated either by the textual content of the visualizations, contextual factors like users' earlier experiences (with visualizations, their subject matter, or other relevant phenomena), or by the physical and psycho-social situation of use. In the analysis of their research findings, Kennedy and Hill cite Jagger (1989) who argues that 'emotion can be understood as an "epistemic resource", a way of knowing that is valuable for building a critique of the world' and Damasio (2006), who argues that without emotions, 'the ability to make rational decisions is hampered' (Kennedy and Hill, 2017, p. 12). Emotions are vital components for understanding the social world, including data visualizations. As such, they are a central focus in this book.

Our emotional engagement with data visualizations is also closely connected to their aesthetic aspects. The forms, colours, and arrangements of data visualizations trigger our senses in particular ways. In turn, the interplay between the semiotic, meaning-making aspects of data visualizations, and the emotions they evoke is closely related to their ability to elicit social engagement. Here, the concept of *engagement* has several layers. It can refer to the actual interaction with a data visualization, being engaged *with it*, or to emotional and practical responses, getting engaged *by it*. It can also refer to broader audience responses, for example the ways in which data visualizations are mobilized to prompt political engagement. These three aspects of engagement are closely interrelated, as can be seen in several chapters in the book.

We understand 'engaging with' visual representations of data to refer to 'the processes of looking, reading, interpreting and thinking that take place when people cast their eyes on data visualisations and try to make sense of them' (Kennedy et al., 2016b). For people who are not experts in data visualizations but who encounter them with growing frequency, engaging

with them is not straightforward. Without the right skills, the ability to participate in data-driven conversations and decision-making will be off limits to certain groups, existing uneven power relations will be reproduced and new, data-based ones will emerge. This has troubling implications for democracy, or for getting *engaged by* the world around us, including by visualized data.

The expansion of data visualization in society therefore requires a new kind of literacy if it is to enable citizens to act in informed and critical ways. It also requires the assessment of data visualization's role in democracy, and the reassessment of democratic theory in light of developments in data visualization. This means asking a range of questions about the relationship between data visualization and democracy. It also means considering the factors in visualization consumption and production processes that affect engagement, which might include factors which extend beyond textual and technical matters, such as class, gender, race, age, location, political outlook, and education of audience members. Some of the contributions in this collection address these issues.

#### Data visualization as discourse

This book is a contribution to multidisciplinary and multifaceted academic conversation concerning the forms, uses, and roles of data visualization in society. As a collection of chapters which study the conditions under which visualizations are generated, disseminated, and thought to benefit processes of learning, development, and participation, to reuse our own phrase from above, it belongs to the large and diverse field of discourse studies. Although the individual chapters derive from a range of perspectives, the tradition of discourse studies provides a framework. The book leans on a social semiotic understanding of *discourse*—as the situated application of semiotic resources (such as words and images) by human agents in order to construct and share ideas about the world and to perform social action (or make things happen) (Kress, 2010; van Leeuwen, 2005). The potential meanings carried by semiotic resources are dependent on both cultural conventions and the particular situations of use, including the background and motivations of the human participants, the media used to produce and distribute the messages, and the social practice of which the semiotic material is an integrated part. Discourse studies can offer nuanced analyses of the mediated processes of communication in which data visualizations are situated and also illuminate processes of social struggle and control.

A discourse studies approach combines the micro level with the macro level. It focuses on the relations between the specific structures and forms of the semiotic artefact on the one hand, and the social, technological, and cultural contexts which form it and are formed by it, on the other (Fairclough, 2010; Chouliaraki & Fairclough, 1999; van Leeuwen, 2005). The concept of discourse thus offers a theoretical and methodological framework for analysing data visualization in discrete social practices, like journalism, public information campaigning, or health communication. These relations between the micro and the macro, between texts and contexts, are apparent in all chapters of the book, although some focus more on the micro level, and others more on the macro level.

Discourse studies include a range of approaches, from those based on an analysis of how meanings are shaped and negotiated in specific social situations, to critical investigations of how words and images play a role in creating or opposing power structures and social inequalities. The latter approaches are often grouped under the term *critical discourse studies* (or CDA), which was originally theoretically and methodologically modelled by Norman Fairclough (2010). In several chapters in this book, similar critical approaches to the relationship between semiotic practices and social inequalities are used, although the authors do not necessarily all see themselves as discourse studies scholars. Rather, authors adopt such approaches from within a diverse range of disciplines, including gender studies, science and technology studies, (digital) media studies, critical cartography, design, art history, literacy studies, ICT, and the emerging field of data studies. Together, the chapters shine a spotlight on data visualization as an important instance of *text-in-society*.

#### How the book is organized and targeted

The book is organized into five sections. The first, called 'Framing Data Visualization', does the work of framing the contributions in the rest of the book, drawing on a range of conceptual and theoretical resources. The three chapters in this section sketch out three significant issues with which subsequent chapters engage: epistemology, semiotics, and politics respectively. In the first chapter in this section, 'Ways of knowing with data visualization', Jill Walker Rettberg explores the ways of knowing that have historically been privileged by different systems for gathering and visualizing data. Giorgia Aiello then maps out how the strategies deployed in a social semiotic approach can help us to understand data visualization

*in society* in 'Inventorizing, situating, transforming: Social semiotics and data visualization'. In the final chapter in this section, Torgeir Nærland maps out perspectives from which we might approach analyses of data visualization's politics, in 'The political significance of data visualization: Four key perspectives'.

The second section of the book, 'Living and Working with Data Visualization', includes chapters which reflect on diverse experiences of and with data visualization in private and professional settings. In Chapter 5, 'Rain on your radar: Engaging with weather data visualizations as part of everyday routines', Eef Masson and Karin van Es explore uses and evaluations of uses of weather data visualizations in everyday life. This is followed by a chapter by Salla-Maaria Laaksonen and Juho Pääkkönen, which shifts the focus to working environments, and explores the uses of data visualizations in social media analytics companies, their role in knowledge claims, and the mechanisms by which they achieve credibility. The chapter is called 'Between automation and interpretation: Using data visualization in social media analytics companies'. Chapter 7, 'Accessibility of data visualizations: An overview of European statistics institutes', by Mikael Snaprud and Andrea Velazquez, uses multiple approaches to assess the extent to which dataviz shared by National Statistics Institutes (NSIs) are accessible to people with disabilities, and the extent of preparedness for compliance with new EU legislation on web accessibility of NSIs, which are both important characteristics of democratic societies. This is followed by a chapter which explores how data visualizations are evaluated, and whether approaches to evaluation which account for the sociocultural contexts of and influences on dataviz might be possible. This chapter, by Arran Ridley and Christopher Birchall, is called 'Evaluating data visualization: Broadening the measures of success.' The subsequent chapter, 'Approaching data visualizations as interfaces: An empirical demonstration of how data are imag(in)ed', by Daniela van Geenen and Maranke Wieringa focuses on the case of a specific data visualization produced by the authors, to show how visualization practices allow for interfacing with data and that a particular visualization provides only one perspective on data. In Chapter 10, 'Visualizing data: A lived experience', Jill Simpson draws on her own experience of producing a small-data hand-drawn visualization to explore questions of subjectivity, authenticity, and honesty in data visualization. This section ends with a chapter by Helen Kennedy, Wibke Weber, and Martin Engebretsen called 'Data visualization and transparency in the news', which explores the relationship between data visualization and the emerging journalistic norm of transparency.

The third section, 'Data visualization, learning, and literacy' includes four chapters which focus on the skills needed to engage with and make sense of data visualizations. In Chapter 12, 'What is visual-numeric literacy, and how does it work?', Elise Seip Tønnessen reports on research based on observations of Norwegian social science classrooms in which students sought to develop and deploy skills to make sense of visualizations in an educational setting. The setting of the next two chapters moves beyond educational institutions. In Catherine D'Ignazio and Rahul Bhargharva's chapter, 'Data visualization literacy: A feminist starting point', the authors introduce a starting point for teaching data visualization which is grounded in feminist theory, process, and design principles, to counter the problem of unequal human relations produced through data. This is followed by 'Is literacy what we need in an unequal data society?', in which Lulu Pinney unearths the different notions of power that are embedded in different uses of literacy across academic literature, policy, and practice in order to critically interrogate the usefulness of literacy as a term and concept. In the final chapter in this section, 'Multimodal academic argument in data visualization', Arlene Archer and Travis Noakes investigate students' semiotic and rhetorical strategies for making an argument with data visualization and their implications for teaching students to become critical citizens.

The fourth section of the book, called 'Data Visualization Semiotics and Aesthetics', includes contributions which focus on the semiotic, aesthetic, visual, and stylistic dimensions of data visualizations and the ways these intersect with social and cultural considerations. Chapter 16, 'What we talk about when we talk about beautiful data visualizations', by Sara Brinch, presents an analysis of what is regarded as beautiful within the field of data visualization design, and at the same time interrogates 'beautiful' as an ambivalent and contested concept. The next chapter, 'A multimodal perspective on data visualization', by Tuomo Hiippala, examines the multimodality of data visualizations, or how they combine multiple modes of expression, such as written language, photographs, diagrammatic elements, and illustrations. This is followed by a chapter by Wibke Weber, 'Exploring narrativity in data visualization in journalism', which explores how and when data visualizations tell stories and the narrative constituents in data visualization, in order to argue that understanding how data are transformed into visual stories is key to understanding how facts are shaped and communicated in society. Chapter 19, by Jonathan Gray, is called 'The data epic: Visualization practices for narrating life and death at a distance'. The chapter proposes the notion of the 'data epic' to explore the narrative and affective capacities of distance in the context of 'public data culture'. This is followed by a chapter by Verena Lechner which

focuses on a very specific aspect of data visualization form, the line, a graphical element widely used in data visualizations to signal a connection between other visual elements. The chapter, 'What a line can say: Investigating the semiotic potential of the connecting line in data visualization', investigates the semiotic functions that connecting lines can have and how these functions can be related to variations in form. The final chapter in this section, 'Humanizing data through graphic visualization', by Aria Alamalhodaei, Alexandra Alberda, and Anna Feigenbaum, considers how the emergent areas of Graphic Medicine and Graphic Social Science deal with numeric data in ways that humanize data, encouraging empathy and connection in audiences. Data visualization could learn from these unconventional fields, the authors propose.

The contributions to the final section, entitled 'Data Visualization and Inequalities', focus on the political dimensions of the social and cultural embedding of data visualization. Chapter 22, 'Visualizing diversity: Data deficiencies and semiotic strategies', by John P. Wihbey, Sarah J. Jackson, Pedro M. Cruz, and Brooke Foucault Welles, explores the complicated dynamics that are inherent to the practice of data visualization involving issues of race and identity. The chapter focuses on data from the US Census and the profound questions that are raised as visual forms purport to represent groups, and showcases a visualization produced by the authors to address the challenges that they discuss. This is followed by a chapter by Rosemary Lucy Hill, 'What is at stake in data visualization? A feminist critique of the rhetorical power of data visualizations in the media'. This chapter argues that visualizations relating to abortion often tell a narrow story, remove contextual detail and omit questions important to women's health. The final three chapters of this section and of the book focus on maps as particular visualizations of data. Chapter 24, 'The power of visualization choices: Different images of patterns in space', by Britta Ricker, Menno-Jan Kraak, and Yuri Engelhardt, uses a dataset related to the United Nations Gender Inequality Index to demonstrate the numerous decisions that are made in the process of creating a map and the types of representations that result. The next chapter, by Anna Berti Suman, 'Making visible politically masked risks: Inspecting unconventional data visualization of the Southeast Asian haze', investigates the potential of data visualization in stimulating a socially and legally accountable governance of environmental risk affecting public health, focusing on mapping efforts of the Southeast Asian haze performed by environmental NGOs and civil society. Finally, Chapter 26, 'How interactive maps mobilize people in geoactivism', by Miren Gutiérrez, explores how maps are employed in activism to unleash sentiments, focusing on three examples and employing as a lens the emotional turn currently influencing geography.

We structure the book in this way with the aim of highlighting the major issues concerning researchers of data visualization in society; many of the chapters cover more than one of the issues that are named in section titles, of course. The book aims to be accessible to a broad audience interested in data visualization's increasing prominence and visibility and its social role. Chapters are written in an accessible style and are relatively short, including real-world examples. All chapters draw on original academic research, and many of them refer to specific visualization projects and practices. All contribute to academic and public conversation about data visualization in society.

#### References

- Ambrosio, C. (2015). Objectivity and representative practices across artistic and scientific visualization. In: A. Carusi, A. S. Hoel, T. Webmoor, & S. Woolgar (Eds.), *Visualization in the age of computerization* (pp. 118-144). London: Routledge.
- Borgo, R., Kehrer, J., Chung, D. H. S., Maguire, E., Laramee, Robert, S., Hauser, H., Ward, M., & Chen, M. (2013). Glyph-based visualization: Foundations, design guidelines, techniques and applications. *Eurographics State of the Art Reports*, 2013, 39-63. Retrieved from http://diglib.eg.org/EG/DL/conf/EG2013/stars/039-063. pdf

Barnhurst, K. G. (1994). Seeing the newspaper. New York: St. Martin's Press.

- Cairo, A. (2013) *The functional art: An introduction to information graphics and visualization*. Berkeley, CA: New Riders.
- Chouliaraki, L., & Fairclough, N. (1999). *Discourse in late modernity: Rethinking critical discourse analysis.* Edinburgh: Edinburgh University Press.
- Couldry, N. (2016). Foreword. In: S. Kubitschko & A. Kaun (Eds.), *Innovative methods in media and communication research*. (pp. i-viii). Cham: Palgrave Macmillan.
- Damasio, A. R. (2006). *Descartes' error: Emotion, rationality and the human brain*. London: Vintage.
- Engebretsen, M., Kennedy, H., Weber, W. (2017). Visualization practices in Scandinavian newsrooms: A qualitative study. *21st International Conference Information Visualisation (IV)*, 297-300. http://doi.org/10.1109/iV.2017.54).
- Fairclough, N. (2010). *Critical discourse analysis: The critical study of language*. London: Routledge.
- Few, S. (2008, August). What ordinary people need most from information visualization today. *Perceptual Edge: Visual Business Intelligence Newsletter*. Retrieved from http://www.perceptualedge.com/articles/visual\_business\_intelligence/ what\_people\_need\_from\_infovis.pdf

- Friendly, M. (2008). A brief history of data visualization. In: C.-H. Chen, W. Härdle, & A. Unwin (Eds.), *Handbook of data visualization*. (pp. 15-56). Berlin: Springer.
- Halliday, M. A. K. (1978). Language as social semiotic: The social interpretation of language and meaning. London: Arnold.
- Jaggar, A. M. (1989). Love and knowledge: Emotion in feminist epistemology. *Inquiry*, 32(2), 151-176. https://doi.org/10.1080/00201748908602185
- Kennedy, H., & Allen, W. (2016). Data visualisation as an emerging tool for online research. In: N. G. Fielding, R. M. Lee, & G. Blank (Eds.), *The Sage handbook of online research methods* (2nd ed.). (pp. 307-326). London: Sage.
- Kennedy, H., & Hill, R. L. (2017). The feeling of numbers: Emotions in everyday engagements with data and their visualisation. *Sociology*, *5*2(4), 830-848. https://doi.org/10.1177/0038038516674675
- Kennedy, H., Hill, R. L., Allen, W., & Kirk, A. (2016a). Engaging with (big) data visualizations: Factors that affect engagement and resulting new definitions of effectiveness. *First Monday*, 21(11). https://doi.org/10.5210/fm.v2111.6389
- Kennedy, H., Hill, R. L., Aiello, G., & Allen, W. (2016b). The work that visualisation conventions do. *Information, Communication and Society*, 19(6), 715-735. https:// doi.org/10.1080/1369118X.2016.1153126
- Kirk, A. (2016). Data visualisation: A handbook for data driven design. London: Sage.
- Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. London: Sage Publications.
- Kress, G., & van Leeuwen, T. (1996) *Reading images: The grammar of visual design*. London: Routledge.
- Latour, B. (1986). Visualization and cognition: Drawing things together. *Knowledge and Society, 6*, 1-40.
- Lemke, J. (2015). Feeling and meaning: A unitary bio-semiotic account. In: P. P. Trifonas (Ed.), *International handbook of semiotics*. (pp. 589-616). New York & London: Springer.
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think.* Boston, MA: Houghton Mifflin Harcourt.
- Porter, T. M. (1995). *Trust in numbers: The pursuit of objectivity in science and public life.* Princeton: Princeton University Press.
- van Leeuwen, T. (2005). Introducing social semiotics. London & New York: Routledge.

#### About the authors

**Helen Kennedy** is Professor of Digital Society at the University of Sheffield. Her research traverses digital landscapes and is currently focused on datafication in everyday life. Her Seeing Data (seeingdata.org) research

into how non-experts relate to data visualizations provides the inspiration for many contributions in this book.

**Martin Engebretsen** is Professor of Language and Communication at the University of Agder, Norway, and director of the INDVIL project (indvil. org), which also provides the inspiration for this book. His research areas include text and discourse studies, multimodality, digital journalism and visual communication.